

Orthogonal polynomials—Constructive theory and applications *

Walter GAUTSCHI

Department of Computer Sciences, Purdue University, West Lafayette, IN 47907, U.S.A.

Received 20 August, 1984

Abstract:

We consider polynomials orthogonal with respect to some measure on the real line. A basic problem in the constructive theory of such polynomials is the determination of their three-term recurrence relation, given the measure in question. Depending on what is known about this measure, there are different ways to proceed. If, as is typical in applications in the physical sciences, one knows the measure only through moment information, the appropriate procedure is an algorithm that goes back to Chebyshev. The algorithm in effect implements the nonlinear map from the given moments (or modified moments) to the desired recursion coefficients. The numerical effectiveness of this procedure is determined in an essential way by the numerical condition of this map. It is known that the map is ill-conditioned in the case of ordinary moments. For modified moments, it may or may not be well-conditioned, the matter depending on certain properties of the given measure and the additional measure defining modified moments. A theorem to this effect will be given and illustrated on some typical examples.

If more is known about the given measure, for example, if it is absolutely continuous and can be evaluated pointwise, then a procedure can be employed which is based on an observation of Stieltjes. Stieltjes remarked that the desired recursion coefficients can be successively built up by alternating between known inner product formulae for these coefficients and the initial sections of the recurrence relation already obtained. An effective implementation of this idea requires a suitable discretization of the inner product. This requirement, while possibly a weakness of the method, also accounts for its beauty, since it leaves room for imagination and ingenuity. Used skillfully, the discretized Stieltjes procedure is among the most widely applicable and effective methods for generating orthogonal polynomials. Some examples will be given to illustrate its use.

Finally, we show how our newly acquired capability of generating nonstandard orthogonal polynomials can be used to solve some special problems in approximation theory and in the summation of slowly convergent series. A novel set of polynomials orthogonal on the semicircle will also be mentioned briefly in connection with Cauchy principal value integrals.

1. Introduction

Classical orthogonal polynomials are being widely used in many branches of science: theoretical physics, chemistry, applied mathematics, probability, approximation theory, numerical analysis, and others. They are easily generated by recursion, and their use is supported by a highly developed analytical theory. Orthogonal polynomials relative to general, nonclassical weight functions, in contrast, have not found the same widespread use, partly because they are more

* This work was supported in part by the National Science Foundation under grant DCR-8320561.

difficult to generate numerically, and also, undoubtedly, because they do not enjoy the close ties with the fundamental differential equations of mathematical physics, as do their classical fellow polynomials. It is nevertheless our belief that, once the constructive problems related to general orthogonal polynomials are solved, nonstandard applications of orthogonal polynomials will be greatly encouraged and the use of nonclassical orthogonal polynomials will become more pervasive than is presently the case.

In the following we give a survey of recent work on the constructive theory of (general) orthogonal polynomials and also indicate some new applications. We begin in Section 2 with formulating what we consider to be the central problem: the generation of the three-term recurrence relation. An important algorithm for this purpose, due essentially to Chebyshev, will be discussed in Section 3. The numerical properties of this algorithm depend largely on the conditioning of the underlying nonlinear map, which is the subject of Section 4. An alternative (and often more effective) constructive approach, which in its key ideas goes back to Stieltjes, will be discussed in Section 5. The final Section 6 contains new applications, all involving, in one form or another, Gauss–Christoffel quadrature.

Our interest here is with polynomials orthogonal on the real line. Polynomials orthogonal on the circle (Szegő's polynomials), or on more general curves or domains in the complex plane, as well as orthogonal polynomials in several variables, are beyond the scope of this survey. They present their own numerical problems which are still largely unexplored. In one of our applications of Section 6, however, we briefly mention certain polynomials orthogonal (but not positive definite) on the semicircle.

2. A basic problem

We are given a positive measure $d\sigma(t)$ on the real line \mathbb{R} , where $\sigma(t)$ is assumed to have at least $n + 1$ points of increase and the first $2n$ moments

$$\mu_k = \int_{\mathbb{R}} t^k d\sigma(t), \quad k = 0, 1, 2, \dots, 2n - 1, \quad (2.1)$$

are finite, with $\mu_0 > 0$. There exists, then, a unique set of (monic) orthogonal polynomials $\pi_k(\cdot) = \pi_k(\cdot; d\sigma)$, $k = 0, 1, \dots, n$, defined by

$$\begin{aligned} \pi_k(t) &= t^k + \text{lower degree terms,} \\ \int_{\mathbb{R}} \pi_k(t) \pi_l(t) d\sigma(t) &\begin{cases} = 0 & \text{if } k \neq l, \\ > 0 & \text{if } k = l, \end{cases} \quad 0 \leq k, l \leq n. \end{aligned} \quad (2.2)$$

They satisfy, as is well-known, a three-term recurrence relation

$$\begin{aligned} \pi_{k+1}(t) &= (t - \alpha_k) \pi_k(t) - \beta_k \pi_{k-1}(t), \quad k = 0, 1, \dots, n - 1, \\ \pi_{-1}(t) &= 0, \quad \pi_0(t) = 1, \end{aligned} \quad (2.3)$$

where $\alpha_k = \alpha_k(d\sigma)$ are real constants, and $\beta_k = \beta_k(d\sigma) > 0$.

A fundamental problem in the constructive theory of orthogonal polynomials is the following:

$$\begin{aligned} &\text{Given } d\sigma(t) \text{ on } \mathbb{R}, \text{ and } n, \text{ compute} \\ &\alpha_k(d\sigma), \beta_k(d\sigma) \text{ for } k = 0, 1, \dots, n - 1. \end{aligned} \quad (2.4)$$

(While β_0 is unimportant in (2.3), it is convenient in other contexts to define, as we do here, $\beta_0 = \int_{\mathbb{R}} d\sigma(t)$.) Another problem is to compute the zeros of $\pi_n(\cdot; d\sigma)$ and additional quantities related to π_n , such as the Christoffel numbers. We will see in section 6.1, however, that this problem can easily be solved, once we have dealt with the problem (2.4). Knowledge of the coefficients $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$ (not only for $k < n$, but also for k sufficiently beyond n) also permits us to compute the functions

$$\rho_r(z) = \int_{\mathbb{R}} \frac{\pi_r(t; d\sigma)}{z - t} d\sigma(t), \quad r = 0, 1, \dots, n, \quad (2.5)$$

where z is complex outside the support of $d\sigma$, in particular Stieltjes' integral $\rho_0(z)$, as minimal solution of the recurrence relation (2.3) (where t is to be replaced by z), using the starting value $\rho_{-1}(z) = 1$ (see Gautschi [11]). This in turn has applications in error analysis of Gaussian quadrature for analytic functions (Gautschi and Varga [23]). Equally important, but more difficult, is the inverse problem of (2.4): *Given the coefficients α_k , β_k , for all $k \geq 0$, and additional asymptotic information for $k \rightarrow \infty$, determine the measure $d\sigma$, including the structure of its spectrum.* There seems to be little numerical experience on this problem, and we shall not consider it here any further. For important analytical techniques, however, see Askey and Ismail [1].

Introducing the vector of recurrence coefficients

$$\rho = [\alpha_0, \dots, \alpha_{n-1}, \beta_0, \dots, \beta_{n-1}]^T \in \mathbb{R}^{2n}, \quad (2.6)$$

Problem (2.4), in effect, requires us to carry out the map $\Sigma(\mathbb{R}) \rightarrow \mathbb{R}^{2n}$ defined by $d\sigma \rightarrow \rho$, where $\Sigma(\mathbb{R})$ is some appropriate measure space on \mathbb{R} . Such a map, of course, cannot be handled on a computer, since $\Sigma(\mathbb{R})$ in general is infinite-dimensional. To make the problem manageable, we must reduce it to finite dimension. This can be done, for example, by departing not from $d\sigma$, but from its first $2n$ moments (2.1). We are assuming, more generally, that we are given the first $2n$ modified moments of $d\sigma$,

$$m_k = \int_{\mathbb{R}} p_k(t) d\sigma(t), \quad k = 0, 1, 2, \dots, 2n-1, \quad (2.7)$$

where $\{p_k\}$ is a given system of polynomials satisfying

$$\begin{aligned} p_{k+1}(t) &= (t - a_k)p_k(t) - b_k p_{k-1}(t), \\ p_{-1}(t) &= 0, \quad p_0(t) = 1, \end{aligned} \quad k = 0, 1, 2, \dots, 2n-2, \quad (2.8)$$

and the constants a_k , b_k are known. The problem (2.4) then becomes:

$$\begin{aligned} &\text{Given the first } 2n \text{ modified moments (2.7) of } d\sigma, \\ &\text{compute } \alpha_k(d\sigma), \beta_k(d\sigma) \text{ for } k = 0, 1, \dots, n-1. \end{aligned} \quad (2.4')$$

With m denoting the vector of modified moments,

$$m = [m_0, m_1, \dots, m_{2n-1}]^T \in \mathbb{R}^{2n}, \quad (2.9)$$

and ρ being as in (2.6), we now have the finite-dimensional (nonlinear) map

$$K_n: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n} \quad m \rightarrow \rho. \quad (2.10)$$

The map K_n is classical in the case where $a_k = b_k = 0$ in (2.8), i.e., $p_k(t) = t^k$, hence $m_k = \mu_k$; it can be expressed explicitly in determinantal form. Unfortunately, this map becomes highly

ill-conditioned, when n is large, and is therefore of little practical use; see, however, Gautschi [14] for an application to the validation of Gaussian quadrature rules. In the next section we present an algorithmic implementation of the map K_n , and in Section 4 report on some results concerning the conditioning of K_n in the case where the polynomials $\{p_k\}$ in (2.8) are themselves orthogonal.

3. Modified Chebyshev algorithm

Chebyshev [5] in 1859 already considered and implemented the map K_n in the special case of ordinary moments, $m = \mu$, and in the case of discrete point measures $d\sigma$. The use of modified moments was first proposed by Sack and Donovan [34] and taken up, among others, by Wheeler [39] to deal effectively with a number of problems in theoretical chemistry involving unknown densities. The algorithm to be presented below is due to Wheeler and generalizes directly the original algorithm of Chebyshev. Since it departs from modified moments, we call it the *modified Chebyshev algorithm*.

The algorithm, basically, obtains the desired recursion coefficients $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$ in terms of the 'mixed moments'

$$\sigma_{k,l} = \int_{\mathbf{R}} \pi_k(t) p_l(t) d\sigma(t), \quad k \leq l, \quad (3.1)$$

which are continually updated as the process unfolds. Note, first of all, that $\sigma_{k,l} = 0$ if $k > l$, by orthogonality. Next we have the obvious relations

$$\begin{aligned} \sigma_{-1,l} &= 0, & l &= 1, 2, \dots, 2n-2, \\ \sigma_{0,l} &= m_l, & l &= 0, 1, \dots, 2n-1, \\ \alpha_0 &= a_0 + m_1/m_0, & \beta_0 &= m_0, \end{aligned} \quad (3.2^0)$$

which serve to initialize the algorithm. The updating of the $\sigma_{k,l}$ and the computation of the α_k , β_k then proceeds according to the formulae

$$\left. \begin{aligned} \sigma_{k,l} &= \sigma_{k-1,l+1} - (\alpha_{k-1} - a_l) \sigma_{k-1,l} - \beta_{k-1} \sigma_{k-2,l} + b_l \sigma_{k-1,l-1}, \\ & \quad l = k, k+1, \dots, 2n-k-1, \\ \alpha_k &= a_k + \frac{\sigma_{k,k+1}}{\sigma_{k,k}} - \frac{\sigma_{k-1,k}}{\sigma_{k-1,k-1}}, \\ \beta_k &= \frac{\sigma_{k,k}}{\sigma_{k-1,k-1}}. \end{aligned} \right\} \quad k = 1, 2, \dots, n-1. \quad (3.2)$$

These can easily be derived from the facts $\sigma_{k+1,k-1} = 0$, $\sigma_{k+1,k} = 0$, and from the two recurrence relations (2.3) and (2.8). (For details, see, e.g., [16, Section 5.4]). The complexity of the algorithm (3.2) is clearly $O(n^2)$. A schematic representation of the algorithm is given in Fig. 3.1, where the 'star' indicates which quantities $\sigma_{k,l}$ (represented by black dots) are involved in the recursion of (3.2), the one computed at each step being circled. The entries in boxes are used to compute $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$. Those on the diagonal also furnish the normalization constants $\sigma_{k,k} = \int_{\mathbf{R}} \pi_k^2(t) d\sigma(t)$.

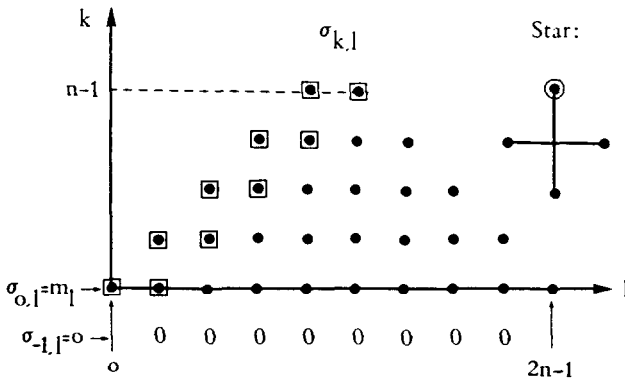


Fig. 3.1. The modified Chebyshev algorithm (schematically for $n = 5$)

The major difficulty with the modified Chebyshev algorithm is the accurate calculation of the modified moments (2.7). In many cases, however, they can be computed explicitly, or by recursion, or can be approximated by a suitable discretization; see Gautschi [16, Section 5.4] for references, and Gautschi [13, Section 4] for a number of examples.

4. Condition of the problem

We have already observed that the map K_n is ill-conditioned in the case of ordinary moments. For measures $d\sigma$ supported on the interval $[0,1]$, for example, the condition of K_n will typically grow like $(3 + \sqrt{8})^{2n}/64n^2$ as $n \rightarrow \infty$, which, incidentally, is the same exponential growth as the one exhibited by the (Turing) condition number of the $n \times n$ Hilbert matrix (Gautschi [8]). On the other hand, experience with the modified Chebyshev algorithm has shown that the use of modified moments does significantly enhance the stability of the algorithm in some cases, but remains ineffective in others. In this section we explore the various factors that are responsible for the conditioning of the map K_n and hence for the stability properties of the modified Chebyshev algorithm.

The basic assumption in this section is that the polynomials p_k defining the modified moments are themselves orthogonal, relative to some measure ds that can be chosen arbitrarily, as appropriate or convenient,

$$\int_{\mathbf{R}} p_k(t) p_l(t) ds(t) = 0, \quad k \neq l. \quad (4.1)$$

What will be interesting to observe is the interplay of the two measures $d\sigma$ and ds in their contribution to the condition of the map K_n .

We first need a convenient representation of the map K_n . It turns out that the Gauss–Christoffel n -point quadrature rule is a good intermediate link; this is the quadrature formula

$$\int_{\mathbf{R}} f(t) d\sigma(t) = \sum_{\nu=1}^n \sigma_{\nu} f(\tau_{\nu}) + R_n(f) \quad (4.2)$$

uniquely determined by the requirement that $R_n(f) = 0$ for all $f \in \mathbb{P}_{2n-1}$. The nodes $\tau_{\nu} = \tau_{\nu}^{(n)}$,

indeed, are the zeros of $\pi_n(\cdot; d\sigma)$, and the weights $\sigma_\nu = \sigma_\nu^{(n)}$ the so-called *Christoffel numbers* for $d\sigma$; see also section 6.1. We introduce the Gauss–Christoffel vector

$$\gamma = [\sigma_1, \dots, \sigma_n, \tau_1, \dots, \tau_n]^T \in \mathbb{R}^{2n}, \quad (4.3)$$

and represent the map $K_n: m \rightarrow \rho$ as a composition of two maps,

$$K_n = H_n \circ G_n, \quad (4.4)$$

where

$$G_n: m \rightarrow \gamma, \quad H_n: \gamma \rightarrow \rho. \quad (4.5)$$

It has been our experience that the map H_n is usually quite well-conditioned, though on rare occasions it, too, may become ill-conditioned; see, e.g., [13, Section 3.1 and 3.4]. We restrict ourselves here to considering the more critical map $G_n: m \rightarrow \gamma$, or, what is technically more satisfying, the map

$$\tilde{G}_n: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n} \quad \tilde{m} \rightarrow \gamma \quad (4.6)$$

from the vector \tilde{m} of *normalized modified moments*

$$\tilde{m}_k = d_k^{-1/2} m_k, \quad d_k = \int_{\mathbb{R}} p_k^2(t) ds(t), \quad k = 0, 1, \dots, 2n-1, \quad (4.7)$$

to the Gauss–Christoffel vector γ . For simplicity, we take the Fréchet derivative of \tilde{G}_n at \tilde{m} ,

$$\tilde{G}'_n(\tilde{m}) \in L(\mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}) \quad (4.8)$$

—a linear bounded operator from \mathbb{R}^{2n} onto itself—or rather a norm of it, as a measure of the sensitivity of the map \tilde{G}_n at \tilde{m} . The condition number of \tilde{G}_n at \tilde{m} (suitably defined) in fact satisfies

$$(\text{cond } \tilde{G}_n)(\tilde{m}) \leq \frac{\|\tilde{m}\|_2}{\|\gamma\|_2} \|\tilde{G}'_n(\tilde{m})\|_F, \quad (4.9)$$

where $\|\cdot\|_2$ is the Euclidean vector norm, and $\|\cdot\|_F$ the Frobenius norm [13, Section 3].

The Frobenius norm of $\tilde{G}'_n(\tilde{m})$ can be expressed exactly in terms of the elementary Hermite interpolation polynomials h_ν and k_ν associated with the Gaussian nodes $\tau_1, \tau_2, \dots, \tau_n$ in (4.2). h_ν is the polynomial in \mathbb{P}_{2n-1} vanishing at all the nodes, except τ_ν , where it has the value 1, and having zero derivative at all n nodes. Similarly, $k_\nu \in \mathbb{P}_{2n-1}$ vanishes at all n nodes and has zero derivative there, except at τ_ν , where it has slope 1. The basic result then is [13, Theorem 3.1]

$$\|\tilde{G}'_n(\tilde{m})\|_F = \left\{ \int_{\mathbb{R}} \sum_{\nu=1}^n [h_\nu^2(t) + \sigma_\nu^{-2} k_\nu^2(t)] ds(t) \right\}^{1/2}, \quad (4.10)$$

where σ_ν are the Christoffel numbers in (4.2).

The critical function in (4.10) is obviously the polynomial

$$g_n(t) = \sum_{\nu=1}^n [h_\nu^2(t) + \sigma_\nu^{-2} k_\nu^2(t)] \quad (4.11)$$

(of degree $4n-2$). It is positive for all real t , and satisfies, as is easily seen,

$$g_n(\tau_\nu) = 1, \quad g'_n(\tau_\nu) = 0, \quad \nu = 1, 2, \dots, n. \quad (4.12)$$

(These relations, of course, do not yet determine g_n .) It is important to recognize that g_n depends solely on the measure $d\sigma$ (through the Gaussian nodes and Christoffel numbers of $d\sigma$). The influence of $d\sigma$ upon the condition of \tilde{G}_n is therefore largely determined by the properties of the polynomial g_n . The second measure ds acts as an integration measure in (4.10) and contributes to the condition of \tilde{G}_n in just that way. To illustrate the interplay of these two measures and their effects on the condition of \tilde{G}_n , we give three examples here; more can be found in [16, Section 5.5].

Example 4.1. $d\sigma(t) = [(1 - k^2 t^2)(1 - t^2)]^{-1/2} dt$ on $[-1, 1]$, $0 < k < 1$; $ds(t) = (1 - t^2)^{-1/2} dt$ on $[-1, 1]$.

The orthogonal polynomials $\pi_r(\cdot; d\sigma)$ in this example are not known explicitly; they have already been used as an example by Christoffel [6]. The polynomials $p_r(\cdot; ds)$ are Chebyshev polynomials, the modified moments therefore Chebyshev moments. These can be computed accurately by recursion [13, Example 4.4]. The polynomial g_n , for many values of k , even close to 1, was found by computation to be ≤ 1 on $[-1, 1]$, without exception. Assuming this to be true, one obtains from (4.10)

$$\|\tilde{G}'_n(\tilde{m})\|_F = \left\{ \int_{-1}^1 g_n(t) ds(t) \right\}^{1/2} < \sqrt{\pi},$$

i.e., the Fréchet derivative of \tilde{G}_n is bounded in norm, uniformly for all n . As a consequence, the map \tilde{G}_n is extremely well-conditioned, even for large values of n . This has been impressively confirmed by numerical experiments.

Example 4.2. $d\sigma(t) = t^\alpha \ln(1/t) dt$ on $[0, 1]$, $\alpha > -1$; $ds(t) = dt$ on $[0, 1]$.

Here we have a branch point (unless α is an integer) and a logarithmic singularity both at the origin. The polynomials $p_k(\cdot; ds)$ are the shifted (monic) Legendre polynomials and the associated modified moments are known in closed form. If α is not an integer, for example, one finds (Gautschi [9])

$$\tilde{m}_k = \frac{\sqrt{2k+1}}{\alpha+1} \left\{ \frac{1}{\alpha+1} + \sum_{l=1}^k \left(\frac{1}{\alpha+1+l} - \frac{1}{\alpha+1-l} \right) \right\} \prod_{l=1}^k \frac{\alpha+1+l}{\alpha+1-l}, \quad k = 0, 1, 2, \dots$$

The polynomial g_n here has a rather peculiar, but not atypical behavior, which we describe for the case $\alpha = -\frac{1}{2}$. On about two-thirds through the interval $[0, 1]$, g_n wiggles around, always staying ≤ 1 ; after that, it starts developing spikes of increasing magnitude, reaching a global maximum at $t = 1$ after a final upward surge. For $n = 40$, for example, the largest spike has magnitude 2×10^4 , while $g_n(1) = 4 \times 10^7$. Fortunately, integration in (4.10) has a smoothing effect, and in spite of the violent oscillations, the condition number of $\tilde{G}_n(\tilde{m})$, as estimated by (4.9), turns out to be relatively small (as condition numbers go). For example, $(\text{cond } \tilde{G}_n)(\tilde{m})$ is less than 3×10^2 , 7×10^2 , 2×10^3 for $n = 20, 40, 80$, respectively (always for $\alpha = -\frac{1}{2}$). The modified Chebyshev algorithm indeed furnishes results very close to machine precision.

Example 4.3.

$$d\sigma(t) = \begin{cases} \pi^{-1} |t - \frac{1}{2}| \{t(1-t)(\frac{1}{3}-t)(\frac{2}{3}-t)\}^{-1/2} dt, & t \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \\ 0 & \text{elsewhere.} \end{cases}$$

This is a measure of interest in the diatomic linear chain (Wheeler [40]). It is interesting because of its support consisting of two separate intervals. We consider two choices of ds .

(a) $ds(t) = \pi^{-1} \{t(1-t)\}^{-1/2} dt$ on $[0,1]$. Thus, $p_k(\cdot; ds)$ are the shifted (monic) Chebyshev polynomials, and m_k Chebyshev moments. For their accurate calculation, we refer to the cited reference. The Gaussian nodes τ_r of $d\sigma$ are known to congregate on the two separate support intervals, except for one node (if n is odd) at the midpoint $t = \frac{1}{2}$. As a result, the polynomial g_n is found to wiggle rapidly on the two support intervals, remaining ≤ 1 there, but shoots up to a huge peak (double peak, if n is odd) on the central interval $[\frac{1}{2}, \frac{2}{3}]$. For $n = 40$, for example, the peak value is of the order 10^{20} ! Since the support of ds is the whole interval $[0,1]$, integration in (4.10) goes right through the peak and produces a large value for the norm of the Fréchet derivative of \tilde{G}_n . For example,

$$\|\tilde{G}'_n(\tilde{m})\|_F = \left\{ \int_0^1 g_n(t) ds(t) \right\}^{1/2} = 2.0 \times 10^9 \quad \text{if } n = 40.$$

Clearly, the 'hole' $[\frac{1}{3}, \frac{2}{3}]$ is to be avoided at all cost. This motivates the second choice of ds .

$$(b) \quad ds(t) = \begin{cases} 18\pi^{-1} |t - \frac{1}{2}|^{-1} \{t(1-t)(t - \frac{1}{3})(t - \frac{2}{3})\}^{1/2} dt, & t \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \\ 0 & \text{elsewhere.} \end{cases}$$

By a stroke of luck, the corresponding modified moments are explicitly computable (Wheeler [40]). The polynomial g_n , of course, being the same as before, we now have

$$\|\tilde{G}'_n\|_F = \left\{ \int_{[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]} g_n(t) ds(t) \right\}^{1/2} < 1, \quad \text{all } n,$$

i.e., the Fréchet derivative is bounded uniformly in n , in striking contrast to the first choice of ds in (a).

It so happens that the recursion coefficients $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$ can be computed in closed form; see Gautschi [18]. This somewhat lessens the practical significance of Example 4.3; nevertheless, the example clearly shows the importance of a proper choice of the measure ds and may provide guidance in similar, but more complicated situations.

5. Discretized Stieltjes procedure

The measure $d\sigma$, in view of (2.1), defines an inner product

$$(p, q) = \int_{\mathbf{R}} p(t) q(t) d\sigma(t) \quad (5.1)$$

on the space \mathbf{P}_{n-1} , or more precisely, for polynomials p and q whose degrees add up to less than $2n$. Stieltjes [37] already observed how the recurrence relation

$$\begin{aligned} \pi_{k+1}(t) &= (t - \alpha_k) \pi_k(t) - \beta_k \pi_{k-1}(t), \quad k = 0, 1, \dots, n-1, \\ \pi_{-1}(t) &= 0, \quad \pi_0(t) = 1, \end{aligned} \quad (5.2)$$

can be combined with the inner product formulae

$$\begin{aligned}\alpha_k &= \frac{(t\pi_k, \pi_k)}{(\pi_k, \pi_k)}, \quad k = 0, 1, \dots, n-1, \\ \beta_0 &= (\pi_0, \pi_0), \quad \beta_k = \frac{(\pi_k, \pi_k)}{(\pi_{k-1}, \pi_{k-1})}, \quad k = 1, \dots, n-1,\end{aligned}\tag{5.3}$$

to gradually build up the recurrence coefficients $\alpha_k = \alpha_k(d\sigma)$, $\beta_k = \beta_k(d\sigma)$, $k = 0, 1, \dots, n-1$. Indeed, since $\pi_0 = 1$, we can use (5.3) to obtain α_0 , β_0 . Applying (5.2) with $k = 0$ then gives π_1 , which together with π_0 allows us to compute α_1 , β_1 from (5.3). Knowing α_1 , β_1 , we can use again (5.2) with $k = 1$ to find π_2 , which together with π_1 gives α_2 , β_2 from (5.3), etc. We call this procedure, alternating between (5.2) and (5.3), the *Stieltjes procedure*. Its main weakness is the necessity of computing the inner products in (5.3). If one tries to do this analytically, by expressing them in terms of the moments (2.1) and the coefficients of the π_k , one in effect implements the ill-conditioned map $K_n: \mu \rightarrow \rho$ and must expect severe instability; see Section 4. For this reason, we proposed in 1968 (Gautschi [8]) to compute the required integrals by numerical quadrature, ignoring essentially the possible presence of singularities in $d\sigma(t)$. This is equivalent to replacing $d\sigma$ by a discrete N -point measure $d\sigma_N$ (usually with $N \gg n$) and then applying Stieltjes' procedure to the inner product defined by $d\sigma_N$. We therefore call this the *discretized Stieltjes procedure*. In [8] we proposed Fejér's quadrature rule (the interpolatory quadrature rule associated with Chebyshev nodes) as a vehicle of discretization, and in [13] the composite Fejér rule. We pointed out, however, that these should be considered merely "discretizations for all seasons", and that it is usually far better to adapt the discretization to prevailing special circumstances. In this sense, the discretized Stieltjes procedure requires a good deal of imagination and skill on the part of the user, but when used properly, can be extremely effective. Two examples follow, to illustrate its use.

Example 5.1. $d\sigma(t) = [t/(e^t - 1)]^r dt$ on $[0, \infty]$, $r \in \mathbb{N}$.

Integration with this measure (involving Einstein's function) is of potential use in solid state physics calculations and has also found application in the summation of slowly convergent series (Gautschi and Milovanović [20]); see also section 6.4.

Since for large t the measure behaves like $d\sigma(t) \sim t^r e^{-t}$, it seems natural to do the integration by Gauss-Laguerre quadrature, after substituting a new variable for rt . This leads to the following discretization (p is a polynomial),

$$\begin{aligned}\int_0^\infty p(t) d\sigma(t) &= \frac{1}{r} \int_0^\infty p\left(\frac{t}{r}\right) \left(\frac{t/r}{1 - e^{-t/r}}\right)^r e^{-t} dt \\ &\approx \frac{1}{r} \sum_{k=1}^N \sigma_k^L p\left(\frac{\tau_k^L}{r}\right) \left(\frac{\tau_k^L/r}{1 - e^{-\tau_k^L/r}}\right)^r = \int_0^\infty p(t) d\sigma_N(t),\end{aligned}\tag{5.4}$$

where σ_k^L , τ_k^L are the weights and nodes of the N -point Gauss-Laguerre quadrature rule. The accuracy obtainable in this way, in spite of the poles at $\pm 2\pi i$, is rather satisfactory. For example, the value of N required to obtain the first n recursion coefficients $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$, $k = 0, 1, \dots, n-1$, to 25 correct decimal digits, is $N = 127$ for $n = 12$ and $N = 281$ for $n = 40$, when $r = 1$, and $N = 85$, $N = 201$, respectively, when $r = 2$.

Example 5.2. $d\sigma(t) = t^\alpha K_0(t) dt$ on $[0, \infty]$, $\alpha > -1$, where K_0 is the modified Bessel function.

This measure arises in the asymptotic evaluation of Bessel transforms (Wong [41]). To arrive at a natural discretization, it is useful to recall the behavior of K_0 ,

$$K_0(t) = \begin{cases} R(t) + I_0(t)\ln(1/t), & 0 < t \leq 1, \\ t^{-1/2}e^{-t}S(t), & 1 \leq t < \infty, \end{cases}$$

where R , S , I_0 are smooth functions. The last one is the regular modified Bessel function. For all of them, high accuracy rational approximations are available (Russon and Blair [33]). We write

$$\begin{aligned} \int_0^\infty p(t)t^\alpha K_0(t) dt &= \int_0^1 [R(t)p(t)]t^\alpha dt + \int_0^1 [I_0(t)p(t)]t^\alpha \ln(1/t) dt \\ &\quad + \int_0^\infty [e^{-1}(1+t)^{\alpha-1/2}S(1+t)p(1+t)]e^{-t} dt, \end{aligned}$$

where in the last integral we have shifted the variable so it runs from 0 to ∞ . This now suggests the discretization

$$\int_0^\infty p(t)t^\alpha K_0(t) dt \approx \int_0^1 p(t) d\sigma_N^{(1)}(t) + \int_0^1 p(t) d\sigma_N^{(2)}(t) + \int_0^\infty p(t) d\sigma_N^{(3)},$$

where $d\sigma_N^{(1)}$ involves Gauss–Jacobi quadrature with parameters 0 and α , $d\sigma_N^{(2)}$ Gauss–Christoffel quadrature with measure $t^\alpha \ln(1/t)dt$, and $d\sigma_N^{(3)}$ Gauss–Laguerre quadrature. The first and last are easily generated, as needed, using the method of section 6.1. The second can be generated by the modified Chebyshev algorithm, which we have seen to be quite stable in this case; see Example 4.2. In this way it is possible, for example, when $\alpha = 0$, or $\alpha = -\frac{1}{2}$, to obtain 15 decimal accuracy in the first n recursion coefficients by taking $N = 100$ for $n = 20$, and $N = 160$, for $n = 40$.

6. Applications

There are a number of areas in which orthogonal polynomials have been used extensively. Among the traditional applications we mention orthogonal expansions, least squares approximation, interpolation, and numerical quadrature. Somewhat less known are the connections of orthogonal polynomials with Padé approximation, when the underlying power series has moments as coefficients; see, e.g., Gautschi [15]. More recent applications include those in probability theory, e.g., birth and death processes, Karlin and McGregor [27,28,29,30], Askey and Ismail [1], coding theory, Sloane [35], Sloane and MacWilliams [36], Bannai and Ito [2], Delsarte [7], scattering theory, Case [4], relaxation methods in numerical linear algebra, Stiefel [38], Gautschi and Lynch [19], and in prediction theory and Toeplitz matrix inversion, where Szegő's polynomials on the unit circle are prominently involved, Kailath [26]. Szegő's theorem on the asymptotic behavior of Toeplitz determinants also has important applications in statistical mechanics, McCoy [32].

Here we consider only some recent simple applications to approximation and summation and some (as yet untested) ideas on Cauchy principal value integrals. Since all of these involve Gauss–Christoffel quadrature, we begin with a modern technique of generating the Gaussian nodes and Christoffel numbers from the recursion coefficients of the appropriate orthogonal polynomials.

6.1 Gauss–Christoffel quadrature rules

As is well-known, the n -point Gauss–Christoffel quadrature rule, relative to a positive measure $d\sigma$,

$$\int_{\mathbb{R}} f(t) d\sigma(t) = \sum_{\nu=1}^n \sigma_{\nu}^{(n)} f(\tau_{\nu}^{(n)}) + R_n(f), \quad R_n(\mathbb{P}_{2n-1}) = 0, \quad (6.1)$$

is the interpolatory quadrature rule having as nodes the zeros $\tau_{\nu}^{(n)}$ of the n -th degree orthogonal polynomial $\pi_n(\cdot; d\sigma)$. Assume that the first n recursion coefficients $\alpha_k = \alpha_k(d\sigma)$, $\beta_k = \beta_k(d\sigma)$, $k = 0, 1, \dots, n-1$, for the orthogonal polynomials $\{\pi_k(\cdot; d\sigma)\}$ have already been obtained, for example by the methods of Section 3 and 5, if $d\sigma$ is nonclassical, or from explicit formulae, otherwise. Then we can form the $n \times n$ Jacobi matrix,

$$J_n = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \ddots & \\ & \ddots & \ddots & \sqrt{\beta_{n-1}} \\ 0 & & \sqrt{\beta_{n-1}} & \alpha_{n-1} \end{bmatrix}, \quad (6.2)$$

whose characteristic polynomial is precisely $\pi_n(\cdot; d\sigma)$,

$$\det(tI_n - J_n) = \pi_n(t; d\sigma). \quad (6.3)$$

Therefore, the nodes $\tau_{\nu}^{(n)}$ in (6.1) are just the eigenvalues of J_n . The Christoffel numbers $\sigma_{\nu}^{(n)}$, in turn, can be expressed in terms of the associated normalized eigenvectors u_{ν} ,

$$J_n u_{\nu} = \tau_{\nu}^{(n)} u_{\nu}, \quad u_{\nu}^T u_{\nu} = 1, \quad \nu = 1, 2, \dots, n. \quad (6.4)$$

Indeed,

$$\sigma_{\nu}^{(n)} = m_0 u_{\nu,1}^2, \quad \nu = 1, 2, \dots, n, \quad (6.5)$$

where $u_{\nu,1}$ is the first component of u_{ν} and $m_0 = \int_{\mathbb{R}} d\sigma(t)$ the first moment. Since the eigensystem problem for symmetric tridiagonal matrices can be solved very efficiently by the QR (or QL) algorithm with appropriate shifts, we have in (6.4), (6.5) a convenient and effective method for generating Gauss–Christoffel formulae; see, in this connection, Golub and Welsch [24], Gautschi [10].

6.2 Approximation by step functions

Spherically symmetric functions in \mathbb{R}^d that arise in physics as distributions are sometimes approximated by physicists in terms of piecewise constant functions, matching as many moments as possible. Calder, Laframboise and Stauffer [3], for example, apply this approach to the Maxwell velocity distribution in \mathbb{R}^d , $d \leq 3$. More precisely, let f be a function of the radial distance $r = \|x\|_2$ in \mathbb{R}^d , and assume that

- (i) $f \in C^1(\mathbb{R}_+)$, $f'(r) \leq 0$ on \mathbb{R}_+ ,
- (ii) $\int_0^\infty f(r) r^j dr$, $\int_0^\infty f'(r) r^j dr$ exist for $j = 0, 1, 2, \dots$

The problem then is to determine an approximation

$$f(r) \approx \tilde{f}(r), \quad \tilde{f}(r) = \sum_{\nu=1}^n a_{\nu} H(r_{\nu} - r), \quad (6.6)$$

with H the Heaviside step function $H(t) = 0$, if $t \leq 0$, $H(t) = 1$ if $t > 0$, in such a way that

$$m_j(f) = m_j(\tilde{f}), \quad j = 0, 1, \dots, 2n-1, \quad (6.7)$$

where

$$m_j(g) = \int_0^\infty g(r) r^j dV, \quad (6.8)$$

$dV = [2\pi^{d/2}/\Gamma(\frac{1}{2}d)]r^{d-1}dr$ being the volume element of the spherical shell in \mathbb{R}^d , $d > 1$, and $dV = dr$ if $d = 1$.

A solution is easily obtained in terms of the Gauss–Christoffel formula (6.1) for the (positive) measure

$$d\sigma(t) = -t^d f'(t) dt \quad \text{on } \mathbb{R}_+ \quad (6.9)$$

(Gautschi [17]). Indeed,

$$r_\nu = \tau_\nu^{(n)}, \quad a_\nu = r_\nu^{-d} \sigma_\nu^{(n)}, \quad \nu = 1, 2, \dots, n, \quad (6.10)$$

where $\tau_\nu^{(n)}$ are the Gaussian nodes and $\sigma_\nu^{(n)}$ the corresponding Christoffel numbers for the measure (6.9). For Maxwell's distribution, $f(r) = \pi^{-d/2} \exp(-r^2)$, for example, one finds

$$d\sigma(t) = (2/\pi^{d/2}) t^{d+1} e^{-t^2} dt \quad \text{on } \mathbb{R}_+, \quad (6.11)$$

while for the Bose–Einstein distribution, $f(r) = (e^r - 1)^{-1}$,

$$d\sigma(t) = t^{d-2} [t/(1 - e^{-t})]^2 e^{-t} dt \quad \text{on } \mathbb{R}_+, \quad d \geq 2. \quad (6.12)$$

The powers of t appearing as a factor in both these measures suggest interesting computational problems: How are the recursion coefficients α_k , β_k for a given measure to be modified if the measure is multiplied by a polynomial [a monomial in the case of (6.11) and (6.12)]? Appropriate algorithms for the solution of this problem are developed in Gautschi [12], Golub and Kautsky [25], Kautsky and Golub [31]. In the first reference, the corresponding problem for division by a polynomial is also considered.

6.3 Approximation by splines

Generalizing the problem in (6.6), we now seek a spline of degree m ,

$$\tilde{f}(r) = \sum_{\nu=1}^n a_\nu (r_\nu - r)_+^m, \quad (6.13)$$

approximating f in the same sense (6.7) as before. Under appropriate assumptions on f , analogous to those in (i), (ii) above, one is led to the measure

$$d\sigma(t) = [(-1)^{m+1}/m!] t^{d+m} f^{(m+1)}(t) dt \quad \text{on } \mathbb{R}_+, \quad (6.14)$$

in terms of which the approximant (6.13) is obtained by

$$r_\nu = \tau_\nu^{(n)}, \quad a_\nu = r_\nu^{-d-m} \sigma_\nu^{(n)}, \quad \nu = 1, 2, \dots, n \quad (6.15)$$

(Gautschi and Milovanović [22]). In contrast to (6.9), the measure (6.14) is no longer necessarily positive. For Maxwell's distribution, e.g., one finds

$$d\sigma(t) = (\pi^{-d/2}/m!) t^{d+m} H_m(t) e^{-t^2} dt \quad \text{on } \mathbb{R}_+, \quad (6.16)$$

where H_m is the Hermite polynomial of degree m . Note, however, that $d\sigma$ in (6.14) is positive for every m , if f is totally monotone.

6.4 Summation of series

Infinite series whose terms involve a Laplace transform, or its derivative, at integer values can be expressed as weighted integrals over the original function. This suggests applying Gauss–Christoffel quadrature to these integrals in order to sum the series. For simplicity, we consider only one type of such series and refer to Gautschi and Milovanović [20] for others.

Let F be the Laplace transform of f ,

$$F(z) = \int_0^\infty e^{-zt} f(t) dt, \quad \operatorname{Re} z \geq 1. \quad (6.17)$$

Then

$$S = \sum_{k=1}^{\infty} (-1)^{k-1} F(k) = \int_0^\infty f(t) \frac{dt}{e^t + 1}. \quad (6.18)$$

Since by Watson's lemma the Laplace transform $F(k)$ behaves like a (usually small) power of k^{-1} as $k \rightarrow \infty$, the series in (6.18) is slowly convergent. Gauss–Christoffel quadrature applied to the integral on the right, with $d\sigma(t) = (e^t + 1)^{-1} dt$, on the other hand, will converge quite satisfactorily, if f is a smooth function, giving rise to an effective summation procedure. To construct the necessary orthogonal polynomials, one could apply the discretized Stieltjes procedure, using a discretization analogous to the one in (5.4). Because of the poles in (6.18) at $\pm i\pi$, which are twice as close to the real axis as the poles in (5.4), it is better, however, to employ composite Fejér quadrature to do the discretization. Once the recursion coefficients $\alpha_k(d\sigma)$, $\beta_k(d\sigma)$ are obtained, the procedure in section 6.1 then quickly yields the desired Gauss–Christoffel formulae. In cases where f is not smooth and has integrable singularities, the measure $d\sigma$ must be modified if satisfactory convergence of Gauss–Christoffel quadrature is to be maintained.

Example 6.1. $F(z) = z^{-1}e^{-1/z}$, $f(t) = J_0(2\sqrt{t})$.

Here,

$$S = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} e^{-1/k} = 0.1971079 \dots,$$

and f —a Bessel function of order zero—is an entire function. Gauss–Christoffel quadrature in (6.18) should therefore converge rapidly. The n -point formula indeed gives relative errors of 1.8×10^{-2} , 9.7×10^{-7} , 1.1×10^{-17} for $n = 2, 4, 8$, respectively. Direct summation of S , on the other hand, is plainly unfeasible.

Example 6.2. $F(z) = z^{-1}(z+1)^{-1/2}$, $f(t) = \operatorname{erf} \sqrt{t}$.

This example is to illustrate the case of a nonsmooth function f ; since the error function is an odd function, f has a square root singularity at the origin. We therefore write

$$\begin{aligned} S &= \sum_{k=1}^{\infty} (-1)^{k-1} k^{-1} (k+1)^{-1/2} \\ &= \int_0^\infty \frac{\operatorname{erf} \sqrt{t}}{e^t + 1} dt = \int_0^\infty \frac{\operatorname{erf} \sqrt{t}}{\sqrt{t}} \cdot \frac{\sqrt{t}}{e^t + 1} dt \end{aligned}$$

and use Gauss–Christoffel quadrature with $d\sigma(t) = t^{1/2}(e^t + 1)^{-1}dt$. This again results in satisfactory convergence, the relative errors of the n -point formula now being 9.2×10^{-6} , 1.6×10^{-10} , 4.6×10^{-20} for $n = 5, 10, 20$, respectively. The numerical value of S is 0.5197632....

6.5 Cauchy principal value integrals

Let C_ε , $0 < \varepsilon < 1$, be the contour in the complex plane formed by the upper unit semicircle, the line segment from -1 to $-\varepsilon$, the upper semicircle of radius ε and center at 0 , and the line segment from ε to 1 . By Cauchy's theorem we then have

$$\lim_{\varepsilon \downarrow 0} \int_{C_\varepsilon} \frac{f(z)}{z} dz = 0$$

for any function f analytic on the closed upper unit half disc. Consequently,

$$\int_{-1}^1 \frac{f(t)}{t} dt = i \left\{ \pi f(0) - \int_0^\pi f(e^{i\theta}) d\theta \right\}, \quad (6.19)$$

where the integral on the left is a Cauchy principal value integral. For the integral on the right we propose a (complex) Gauss–Christoffel quadrature rule,

$$\int_0^\pi f(e^{i\theta}) d\theta = \sum_{\nu=1}^n \sigma_\nu^{(n)} f(\zeta_\nu^{(n)}) + R_n(f), \quad (6.20)$$

derived in the usual way from orthogonal polynomials $\{\pi_k(z)\}$, here those orthogonal on the semicircle,

$$\int_0^\pi \pi_k(e^{i\theta}) \pi_l(e^{i\theta}) d\theta = 0, \quad k \neq l. \quad (6.21)$$

In particular, the nodes $\zeta_\nu^{(n)}$ are the (complex) zeros of $\pi_n(z)$.

Although the inner product in (6.21) is not positive definite (the second factor is not conjugated!), one can still show that a unique system of monic (complex) orthogonal polynomials exists. They can be expressed, in fact, as a two-term linear (complex) combination of Legendre polynomials (Gautschi and Milovanović [21]). Moreover, the zeros $\zeta_\nu^{(n)}$ are located symmetrically with respect to the imaginary axis, and are all contained in the (open) upper unit half disc ([21]). As of the time of writing, numerical experience with (6.20) is not yet available.

References

- [1] R.A. Askey and Mourad E.H. Ismail, Recurrence relations, continued fractions and orthogonal polynomials, *Mem. Amer. Math. Soc.* **49** (300) (1984) 1–108.
- [2] E. Bannai and P. Ito, *Algebraic Combinatorics I, Association Schemes* (Benjamin Cummings, Menlo Park, CA, 1984).
- [3] A.C. Calder, J.G. Laframboise and A.D. Stauffer, Optimum stepfunction approximation of the Maxwell distribution, unpublished manuscript.
- [4] K.M. Case, Orthogonal polynomials revisited, in: R.A. Askey, Ed., *Theory and Application of Special Functions* (Academic Press, New York, 1975) 289–304.
- [5] P.L. Chebyshev, Sur l'interpolation par la méthode des moindres carrés, *Mém. Acad. Impér. Sci. St. Pétersbourg* (7) **1** (15) (1859) 1–24; and, *Oeuvres I*, 473–498.

- [6] E.B. Christoffel, Sur une classe particulière de fonctions entières et de fractions continues, *Ann. Mat. Pura Appl.* (2) **8** (1877) 1–10; and, *Ges. Math. Abhandlungen II*, 42–50.
- [7] P. Delsarte, An algebraic approach to the association schemes of coding theory, *Philips Research Reports Supplements* **10** (1973).
- [8] W. Gautschi, Construction of Gauss–Christoffel quadrature formulas, *Math. Comp.* **22** (1968) 251–270.
- [9] W. Gautschi, On the preceding paper ‘A Legendre polynomial integral’ by James L. Blue, *Math. Comp.* **33** (1979) 742–743.
- [10] W. Gautschi, On generating Gaussian quadrature rules, in G. Hämmerlin, Ed., *Numerische Integration*. Internat. Ser. Numer. Math. **45** (Birkhäuser, Basel, 1979) 147–154.
- [11] W. Gautschi, Minimal solutions of three-term recurrence relations and orthogonal polynomials, *Math. Comp.* **36** (1981) 547–554.
- [12] W. Gautschi, An algorithmic implementation of the generalized Christoffel theorem, in: G. Hämmerlin, Ed., *Numerical Integration*, Internat. Ser. Numer. Math. **57** (Birkhäuser, Basel, 1982) 89–106.
- [13] W. Gautschi, On generating orthogonal polynomials, *SIAM J. Sci. Statist. Comput.* **3** (1982) 289–317.
- [14] W. Gautschi, How and how not to check Gaussian quadrature formulae, *BIT* **23** (1983) 209–216.
- [15] W. Gautschi, On Padé approximants associated with Hamburger series, *Calcolo* **20** (1983) 111–127.
- [16] W. Gautschi, Questions of numerical condition related to polynomials, in G.H. Golub, Ed., *Studies in Numerical Analysis* (Math. Assoc. America, Washington, DC, 1985) 140–177.
- [17] W. Gautschi, Discrete approximations to spherically symmetric distributions, *Numer. Math.* **44** (1984) 53–60.
- [18] W. Gautschi, On some orthogonal polynomials of interest in theoretical chemistry, *BIT*, to appear.
- [19] W. Gautschi and R.E. Lynch, Error behavior in optimal relaxation methods, *Z. Angew. Math. Phys.* **33** (1982) 24–35.
- [20] W. Gautschi and G.V. Milovanović, Gaussian quadrature involving Einstein and Fermi functions with an application to summation of series, *Math. Comp.*, to appear.
- [21] W. Gautschi and G.V. Milovanović, Polynomials orthogonal on the semicircle, submitted for publication.
- [22] W. Gautschi and G.V. Milovanović, Spline approximations to spherically symmetric distributions, in preparation.
- [23] W. Gautschi and R.S. Varga, Error bounds for Gaussian quadrature of analytic functions, *SIAM J. Numer. Anal.* **20** (1983) 1170–1186.
- [24] G.H. Golub and J.H. Welsch, Calculation of Gauss quadrature rules, *Math. Comp.* **23** (1969) 221–230.
- [25] G.H. Golub and J. Kautsky, Calculation of Gauss quadratures with multiple free and fixed knots, *Numer. Math.* **41** (1983) 147–163.
- [26] T. Kailath, Notes on the Szegő unit circle orthogonal polynomials in least-squares prediction theory, in: R. Askey, Ed., *Gabor Szegő: Collected Papers 1* (Birkhäuser, Boston, 1982) 43–46.
- [27] S. Karlin and J. McGregor, The differential equations of birth-and-death processes, and the Stieltjes moment problem, *Trans. Amer. Math. Soc.* **85** (1957) 489–546.
- [28] S. Karlin and J. McGregor, Linear growth, birth and death processes, *J. Math. Mech.* (now *Indiana Math. J.*) **7** (1958) 643–662.
- [29] S. Karlin and J. McGregor, Many server queueing processes with Poisson input and exponential service times, *Pacific J. Math.* **8** (1958) 87–118.
- [30] S. Karlin and J. McGregor, Random walks, III, *J. Math.* **3** (1959) 66–81.
- [31] J. Kautsky and G.H. Golub, On the calculation of Jacobi matrices, *Linear Algebra Appl.* **52/53** (1983) 439–455.
- [32] B.M. McCoy, Introductory remarks to Szegő’s paper ‘On certain Hermitian forms associated with the Fourier series of a positive function’, in: R. Askey, Ed., *Gabor Szegő: Collected Papers 1* (Birkhäuser, Boston, 1982) 47–52.
- [33] A.E. Russon and J.M. Blair, Rational function minimax approximations for the Bessel functions $K_0(x)$ and $K_1(x)$, Rep. AECL-3461, Atomic Energy of Canada Limited, Chalk River, Ontario, October, 1969.
- [34] R.A. Sack and A.F. Donovan, An algorithm for Gaussian quadrature given modified moments, *Numer. Math.* **18** (1971/72) 465–478.
- [35] N.J.A. Sloane, An introduction to association schemes and coding theory, in: R.A. Askey, Ed., *Theory and Application of Special Functions* (Academic Press, New York, 1975) 225–260.
- [36] N.J.A. Sloane and F.J. MacWilliams, *The Theory of Error-Correcting Codes I, II* (North-Holland, Amsterdam, 1977).
- [37] T.J. Stieltjes, Quelques recherches sur la théorie des quadratures dites mécaniques, *Ann. Sci. Ecole Norm. Paris Sér. 3*, **1** (1884) 409–426; and, *Oeuvres I*, 377–396.

- [38] E.L. Stiefel, Kernel polynomials in linear algebra and their numerical applications, in: *Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues*, App. Math. Ser., Vol. **49**. (National Bur. Standards, Washington, D.C., 1958) 1–22.
- [39] J.C. Wheeler, Modified moments and Gaussian quadrature, *Rocky Mountain J. Math.* **4** (1974) 287–296.
- [40] J.C. Wheeler, Modified moments and continued fraction coefficients for the diatomic linear chain, *J. Chem. Phys.* **80** (1984) 472–476.
- [41] R. Wong, Quadrature formulas for oscillatory integral transforms, *Numer. Math.* **39** (1982) 351–360.